REPORT NO. 510-QR-1

RESEARCH ON MAN-MACHINE RELATIONSHIPS IN PRELAUNCH CHECKOUT OF ADVANCED SPACE VEHICLES

QUARTERLY PROGRESS REPORT NO. 1

GPO PRICE $ _____
OTS PRICE(S) $ _____
Hard copy (HC) 3.00
Microfiche (MF) .75

FACILITY FORM 602
N65-12729
(ACCESSION NUMBER)
70
(PAGES)
CR 59699
(NASA CR OR TMX OR AD NUMBER)
(THRU)
1
(CODE)
(CATEGORY)

PREPARED FOR:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D. C.
CONTRACT NO. NASw-954

DUNLAP AND ASSOCIATES, INC.
DARIEN, CONNECTICUT

510-QR-1
DRD-64-122

RESEARCH ON MAN-MACHINE RELATIONSHIPS IN PRELAUNCH CHECKOUT OF ADVANCED SPACE VEHICLES

Quarterly Progress Report Number 1

(1 June to 1 September 1964)

Prepared for:

National Aeronautics and Space Administration
Code MAT
Washington, D. C.

Contract NASw-954

14 September 1964

Dunlap and Associates, Inc.
Darien, Connecticut

TABLE OF CONTENTS

INTRODUCTION i

TASK I. FACTORS AFFECTING THE DEVELOPMENT OF CONFIDENCE DURING CHECKOUT

A. Introduction I-1
B. Objective and Subjective Confidence I-1
C. Some Problems of Observing Subjective Confidence I-2
D. Indirect Influence of Subjective Factors I-3
E. Plans for Second Quarter I-4

TASK II. OPTIMIZATION OF ASTRONAUT PARTICIPATION IN PRELAUNCH TEST AND CHECKOUT ACTIVITIES

A. Introduction II-1
B. Work Accomplished II-2
C. Plans for Second Quarter II-7

TASK III. MAN-COMPUTER ROLES IN PRELAUNCH FAILURE DIAGNOSIS

A. Introduction III-1
B. General Context III-1
C. Structure of Elemental Checkout Operations III-3
D. Complex Checkout Operations III-5
E. Deviations from Planned Procedures III-7
F. Plans for Next Quarter III-11

TASK IV. FAILURE ANTICIPATION

A. Introduction IV-1
B. Scope of the Problem IV-2
C. Approach IV-3
D. Specific Statistical Techniques IV-4
E. Plans for Second Quarter IV-10

TASK V. FUTURE MAN-MACHINE ROLES IN SEMIAUTOMATED TEST PLANNING AND PROGRAMMING

A. Introduction V-1
B. Task Organization V-1

INTRODUCTION

This report describes the first three months of effort in our current study of the various man-machine interactions involved in the prelaunch test and checkout of space vehicles. In a sense, this program represents a continuation of the research conducted by Dunlap and Associates, Inc., under subcontract to The RAND Corporation as part of Contract NASr-21(08). This earlier contract involved the development of design guidelines and display requirements for prelaunch checkout systems. During the course of that study, we identified several major areas of man-machine interaction which required further analysis. The current program is designed to investigate those areas and other pertinent problems arising from our investigations.

Briefly stated, the areas of significance involve the analysis of man-machine interaction requirements in:

- the development of confidence as one of the underlying objectives of prelaunch test and checkout
- the role of astronauts and capsule observers in prelaunch checkout
- prelaunch failure diagnosis
- failure anticipation
- test planning and programming

Correspondingly, we established separate project tasks to investigate each of these areas. An additional special task was set up to investigate a set of visual phenomena uncovered during our earlier work for The RAND Corporation. The following discussions of progress and plans are presented by task in order to provide a coherent picture of the goals of each task and the conceptual approaches being followed.

The sections devoted to each task differ in the nature and depth of the analysis involved and in the level of detail presented. A good deal of the initial period was spent in identifying appropriate documents and data and their sources, procuring of these documents, and establishing contact with interested NASA personnel both in OMSF and at MSC. Since some of the tasks are essentially follow-on activities from our previous NASA work, while others represent new

areas of study, we were able to start work more quickly on some tasks than others. This resulted in an unequal distribution of effort on the several tasks during the first quarter. The division of effort and manpower across tasks will not necessarily be equal or constant during subsequent reporting periods. Continuing evaluation of available data and of the pay-off potential of the various research areas will serve to re-direct the emphasis of our research program.

The following sections of this report describe each of the five major project tasks, the special task, and significant administrative aspects of our effort during the first quarter.

TASK I. FACTORS AFFECTING THE DEVELOPMENT OF CONFIDENCE DURING CHECKOUT

A. Introduction

Prelaunch test and checkout activities are organized and conducted for one primary purpose -- to provide assurance that all on-board systems are in a state of readiness for launch and that, therefore, the probability of mission success is sufficiently high to warrant such a launch. A testing program can only yield empirical data on the performance of a system under given conditions and in a given period of time. The verdict that a system is "go" must, therefore, consist of an evaluation, based on direct evidence (i.e., test results) that the system is satisfactory and also an inference that this evidence is sufficiently reliable to permit extrapolation to some future time, i.e., the mission.

The transition from evidence to a launch decision must, therefore, be based on a value judgment regarding the adequacy of the checkout process. This judgment, in turn, is dependent on the test engineer's feeling that (1) the test process has indeed collected all of the known critical and relevant evidence regarding system status, and (2) that it has gathered, evaluated, and reported such data in a reliable and undistorted fashion.

What we are thus saying is that the quality of a test engineer's judgment about system readiness is a function of his confidence both in the comprehensiveness and relevance of the test program and in the validty of the test results. The use of the word "confidence" as a concise designation of all of the complex factors which determine when the spacecraft should be accepted as ready for use is extremely convenient. However, the generic meaning of the word does not provide an adequate basis for planning and organizing Task I. For this purpose the concept must be defined in a way which makes it possible to describe the way in which confidence depends upon the various observable factors which might be significant.

B. Objective and Subjective Confidence

Confidence can be described in terms of the weighting a person gives to the evidence resulting from an observation, a datum, or a process. The factors affecting confidence fall into two categories: objective (empirical) and subjective (personal). The weightings assigned on the basis of subjective and objective factors may or may not be highly correlated with each other and may even be largely independent of the actual amount of evidence available. Objective

confidence in test results is primarily a matter of the collection of direct empirical evidence with demonstrated precision, reliability, and repeatability and with sufficient redundancy. Subjective confidence, on the other hand, is a more personal expression of a test engineer's evaluation of the test situation which may or may not be highly correlated with objective evidence. Thus, his past experience and present attitudes may cause him to weigh the evidence in a unique way or to disregard it entirely in some instances.

If only the objective aspects of confidence had to be considered, the problems of developing test and checkout procedures would be relatively uncomplicated. However, experience makes it clear that conformance with all of the objective criteria is a necessary but not sufficient condition for achieving an adequate level of confidence. The additional factors involved are best described as "subjective" and include all of the past experience and present attitudes applicable to a given situation. The dominant influence of these subjective factors is clearly demonstrated by the fact that the final decision to accept or reject equipment is always left to the "judgment" of responsible individuals who have the option of completely ignoring the objective evidence if this appears to be warranted by other considerations. For this reason, identification and description of these other factors is an important task in this study.

It must be recognized at the outset, however, that this study objective is paradoxical: as subjective criteria are identified and described, they tend to become objective. This can be demonstrated by examining some hypothetical examples of situations which, at first sight, appear to provide clear descriptions of the role of subjective factors influencing confidence.

C. Some Problems of Observing Subjective Confidence

If the concept of subjective confidence is to be useful in the study, it must be, in some sense, an observable phenomenon (i.e., its relationships to the various variable aspects of test and checkout situations must be determinable). Since, by definition, subjective confidence is intangible, it cannot be observed directly. Its characteristics must, therefore, be inferred from its influence on overt behavior. In this study, the observable behavior of interest is the acceptance or rejection of equipment (the system(s) being tested) the ultimate acceptance being the decision to launch. We can attempt to visualize situations in which the effect of subjective confidence in arriving at these decisions is unambiguously detectable.

The first possibility is a comparison of two situations in which the objective factors are identical, but where the equipment was accepted in one instance and rejected in the other. Can the difference between these two cases be ascribed

to a difference in subjective confidence in the capabilities of the equipment? One answer to this question is that there is at least one other set of factors, independent of confidence, which could also produce this result; namely, degree of risk. As an example, suppose that the equipment being considered is a booster. It might be accepted for use when the vehicle is unmanned but rejected when a man's life is at stake, with no difference whatsoever in the confidence in the capabilities of the equipment, because the risk of loss in the second case is not acceptable. Estimates of risk loss (and, by extension, chance of gain) are frequently highly subjective. They are sensitive to such intangibles as political pressure, status in the "space race," personal ambitions, military exigency, publicity, personality traits, and budgetary restrictions. The risk factor tends to set the acceptable level of confidence in equipment but does not affect this confidence per se.

These considerations suggest modification of the original example to include equal risk as well as equal subjective confidence. In this case, subjective confidence appears to be a reasonable basis for explaining the difference in action. However, attempts to develop practical instances indicate that the example is one which will never occur in practice. The two cases might start with the indicated identity, and the key factor initiating the difference in results might be completely subjective, e.g., a "hunch." However, a competent, experienced, technical expert would not merely change his opinion at this point and let it go at that. He would search further for some rational basis for selecting a different action in the second instance. If a convincing rationalization cannot be found, the subjective indications will be ignored and there will be no difference between the two cases. If a satisfactory rationalization is found, the objective procedures in the two cases are no longer identical. This example, therefore, produces no useful basis for investigation of the effects of subjective factors.

D. Indirect Influence of Subjective Factors

The way in which subjective factors were used in the preceding example suggests a generalization which provides another approach to the problem of identifying the effects of these factors. The subjective factors are not, per se, a basis for acceptance or rejection of equipment; rather, they have an indirect effect as a result of their influence on the rational objective criteria for acceptance. Specifically, the men responsible for the procedures can do the following:

- They can change the acceptance criteria by specifying higher (or lower) performance levels, smaller (or larger) variations from standards, etc.

- They can introduce additional tests or criteria including rechecking, testing of equipment normally accepted without checking, collection of additional data, etc.

- They can reinterpret test data by noting transients, rates of change, new combinations of data, etc., which were not previously considered significant.

- They can alter operational plans so as to eliminate the need for particular items of equipment, introduce "jury rigs," etc.

- Most important, they can do all this on an ad hoc basis, that is, as an integral part of the checkout operation, in response to the immediate circumstances, without conscious preplanning, and very frequently, without completely rigorous logical justification for each decision.

These are the unique capabilities -- especially the last one -- which make competent human beings essential in operations as well as in planning, and it is the human behavior relating to these functions which is affected primarily by subjective factors. Man's ability to perform these functions can be categorized in terms of the following characteristics:

- Intelligence, including the ability for creative, inductive (or more generally, heuristic) reasoning; the ability to learn.

- Knowledge of the equipment, its performance characteristics, history, etc. based upon training and experience.

- Personality factors such as control of optimistic or pessimistic excesses, motivation, personal dedication to the task and its objectives, etc.

When the question of the significance of subjective factors in checkout operations is characterized in this way, the problem of identifying observable factors can be resolved in a very practical and intuitively satisfying way. We can postulate the existence of a completely planned checkout procedure and of criteria which associate each possible result of implementing this procedure with a terminal decision to accept or reject the equipment. In any particular instance, it is then possible to observe deviations from the accepted procedure or criteria. These deviations can then be analyzed on a post mortem basis to determine:

- The incident, observation, idea, or what-have-you which initiated the observed deviation.

- The rational justification (if any) for modifying the procedure or criteria in the observed manner.

- The desirability of making this modification a part of a new standard procedure or criterion.

In this approach, confidence is indicated by conformance with established procedures and criteria. Conversely, lack of confidence or attempts to improve confidence are indicated by the introduction of ad hoc deviations from these established standards.

In summary, the study can be broken down into the following subtasks:

- Historical and "in vivo" observation of the effect of ad hoc deviations on the evolutionary development of checkout procedures and acceptance criteria.

- Identification of the environmental, psychological, and technological factors which determine the nature and frequency of these ad hoc innovations.

- Evaluation of the benefits or ill effects of these innovations for both the operation in which they were introduced and the subjequent operations.

- Generation of recommendations for general procedures which capitalize on the advantages obtainable from ad hoc innovation while guarding against the possible disadvantages.

The first step in such a study is to develop a description of the postulated "completely planned procedure" so that both the procedure itself and the deviations from it can be identified in practical situations. This description is relevant to both Task I and III and is discussed in the section devoted to the latter task.

E. Plans for Second Quarter

This initial period was devoted to the definition and structuring of this study area. During the subsequent reporting period, the Task I effort will consist of the following activities:

1. Continued exploration of the psychological and operational factors involved in establishing and increasing confidence. This will include discussions and structured interviews with test conductors to determine:

   - The extent of their participation in the development of test procedures and requirements.

- Their current attitudes towards the introduction of automated checkout procedures.

- Their feelings about the amount of test data that will be available to them and the format in which it will be presented.

- Their feelings about the relative reliability of manual vs. automatic checkout.

- Etc.

2. Observations of on-going spacecraft and booster tests to isolate instances of ad hoc deviations from preplanned procedures and the reasons therefor.

3. Continued review of psychological and other literature relating to the establishment of value structures and the impact of situational, organizational, and personal variables upon such values.

4. Exploration of various psychometric techniques for evaluating multidimensional value structures.

TASK II. OPTIMIZATION OF ASTRONAUT PARTICIPATION IN PRELAUNCH TEST AND CHECKOUT ACTIVITIES

A. Introduction

Currently, a major constraint on complete automation of prelaunch test and checkout for the Apollo spacecraft is the requirement for astronauts at their crew stations, to participate in checkout activities. Their activities must be conducted in real time and thereby in large measure set the pace for major segments of checkout operation.

True end-to-end testing requires that system operation be initiated within the crew station and that at least some of the outputs be evaluated at the crew station displays. The most obvious method of accomplishing these initiation and monitoring functions is via the use of an on-board operator. Further, for several reasons, the most obvious operator is the astronaut. However, excessive dependence on the astronauts may result in inefficiencies in both checkout procedures and use of astronaut time.

The purpose of this task is to optimize astronaut participation in prelaunch test and checkout activities, both from a checkout or "engineering" standpoint and from a crew training standpoint.

The following is an outline of the work to be done in accomplishing this end.

- Identification of current purposes and objectives of astronaut participation.
- Verification of these objectives in light of future systems.
- Identification of the functions which the astronaut must perform in system and mission operations.
- Examination of test requirements and procedures to identify the functions which must be performed in the crew station during checkout.
- Investigation of other means of performing these functions during checkout.
- Isolation and categorization of test information-flow requirements involving astronaut participation.

. Development of design principles and operating procedures to facilitate astronaut participation.

. Examination of astronaut/checkout-computer interface design requirements.

During the first quarter our activities have been centered upon: identification of the present requirements for astronaut participation in checkout; identification of astronaut functions and information requirements; and examination of preliminary test requirements. From these we have formulated certain hypotheses which we plan to test and have given thought to possible astronaut/computer interface designs. All of these activities are discussed in further detail in the following sections.

B. Work Accomplished

1. Mercury Analysis

We performed an analysis of Project Mercury postflight reports to identify the stated and implied reasons for astronaut participation in a manual checkout system.

This analysis indicated that there were two basic classes of reasons for astronaut participation in Mercury testing and checkout. These are: (1) training (of both the astronaut and the launch crew), and (2) engineering (the astronaut was an essential part of the system). A more detailed breakdown reveals the following reasons:

a. Training Reasons

. Astronaut Training

- Familiarization with specific configuration of S/C and systems (control and display orientation and location, control feel, needle movements, light intensities).

- Detailed knowledge of the systems and the diagnostic procedures performed by system engineers.

- Familiarization with specific system performance (noise and vibration, flow rates, temperatures).

- Familiarization with checkout and launch procedures (astronauts roll-in countdown, reporting procedures, control procedures, timing).

- Familiarization with equipment and procedures for in-flight scientific experiments.

- Familiarization with launch crew and mission controllers and development of confidence in them.

- Ingress and egress practice.

. Launch Crew and Mission Controllers Training

- Familiarization with idiosyncrasies of individual astronauts (voice quality, reactions, operational timing).

b. Engineering Reasons

. Astronaut Specific Flight Preparation

- Adjustments on personal equipment, (location in seat, strap adjustments, hose adjustments, control settings).

- Preparation and installation of personal checklists for S/C and mission operations and scientific experiments (includes personalized marking and labeling of controls and displays).

. Astronaut Brings System Closer to Flight Configuration

- Individual physical characteristics and personal-equipment/spacecraft interface (oxygen flow rates, cooling temperatures, weight and balance, communication system settings).

- Permits checkout of systems normally controlled by astronaut and effected by astronaut performance.

- Provides on-board sensing and stimulus generation as would be expected in flight.

2. Apollo Test Analysis

We reviewed preliminary Apollo test requirements to identify further "engineering" reasons for astronaut or spacecraft observer (i.e., system test engineer or astronaut types) participation during different categories of prelaunch tests. (The training reasons are expected to be similar to those set forth for the Mercury program.)

The study involved examination of the planned testing to be performed on the Apollo spacecraft to determine, from an engineering point of view, what tests would require the active services of an operator (astronaut or system engineer) or operators. Our analysis assumes that any test requiring the operation of on-board switches, circuit breakers or other controls also requires a human operator to perform these operations. A matrix of all major tests versus individual spacecraft systems was constructed. The major tests included all those to be performed at Cape Kennedy, from combined systems tests on a module (CM, SM, LEM) to mission simulations of the entire spacecraft. The spacecraft systems were subdivided into test-oriented subsystems.

The test description for each subsystem within each major test category was then examined and an estimate made of the number and type of persons, if any, required to handle the on-board control or monitoring operation. This estimate was then inserted into each cell of the matrix. The results of this study indicate that approximately 90% of the tests require the human to perform some on-board function. However, the brief test descriptions examined did not permit the clear discrimination between those functions which were involved only in pretest setup and those which were actually operational functions performed during test runs. Furthermore, with the exception of mission simulation tests, it was not possible to differentiate clearly between those functions which require an astronaut and those which can be performed by any other type of on-board observer.

As a result of structuring the spacecraft testing in this manner, three basic types of tests were identified. These are: Combined Systems Tests (CST), Mission Simulation Tests, and Countdown, either simulated or actual. Most of the other major test sequences are made up of combinations of these three, with minor additions or omissions. Even the numerous tests performed in special facilities (e.g., altitude chamber, static fire facility) include many portions of the three basic types. This structuring of test types proves to be useful in examining several other aspects of this task, as will be shown later in this section.

3. Analyses of Mercury and Apollo Test Procedures

This was done to determine the general categories of astronaut activities and their information and control requirements during testing. Several examples of Mercury test scripts (e.g., simulated flight, FACT) and preliminary Apollo system test procedures were examined in step-by-step detail. Every task required of an on-board operator was identified and listed separately. These were then grouped into system-oriented and mission-oriented tasks and categorized as pretest setup or system (mission) operational

tasks. Classification of the individual tasks in this manner resulted in the following categorization of tasks performed by an on-board operator during test and checkout operations.

- Pretest Setup Tasks (both system- and mission-oriented tasks)
  - Receive instruction and/or program information
  - Transmit instruction to S/C (through switches and controls) according to program
  - Observe system status
  - Report status according to program
- System or Mission Operational Tasks
  - Receive operational instruction and/or program information
  - Transmit instruction to S/C (through switches and controls) according to program
  - Observe system feedback
  - Report system feedback according to program
  - Determine system, test, and mission status
  - Report system and mission status according to program

It should be noted that the underlined words in the above outline indicate the four types of information both transmitted and received by the astronaut and/or spacecraft observer during test and checkout operations. One other type, background or reference information, is also necessary. In the astronaut's case, this is primarily in the form of knowledge gained through training and experience.

4. Conceptual Development

The preceding analyses led to a concept of a dual role for astronauts in the checkout process: (1) as a spacecraft subsystem; and (2) as a checkout system stimulus generator and status monitor. This dual role becomes apparent when the tasks required of an astronaut (spacecraft observer) during the three basic types of Apollo tests are examined.

First, because it is a manned spacecraft, primary and secondary mission control and system operation is the responsibility of a human crew following lift-off. Thus the crew, at least in mission operational modes of testing (i.e., mission simulation tests and final countdown), must be considered an essential subsystem, the performance of which must also be tested. As a subsystem, his condition is monitored directly (medically) and his performance is sensed through other system test points. In this type of test, the up-link checkout computer establishes the spacecraft environment and system conditions through GSE. The mission operations are commanded by the crew through their respective spacecraft systems back to the computer. The computer, in turn, transforms these commands into new inputs to the spacecraft sensors and into test readouts. Thus, the computer acts as a slave to the on-board crew in a mission simulation.

In the second case, that of combined system testing, the on-board crew is not acting as a spacecraft subsystem requiring performance verification, but rather as a stimulus generator and status monitor. He performs these functions according to a program and upon instructions, controlled by the checkout computer. In this sense, man essentially becomes a piece of carry-on test equipment.

Several ideas for accomplishing this second category by methods other than on-board human operators have been briefly discussed. Because these discussions are of such a preliminary nature at this time, and because their applicability is probably beyond Apollo, they will only be mentioned here. These include:

- Switching Matrix - A matrix of diodes incorporated into the spacecraft systems permitting the opening and closing of circuits from various sources (e.g., crew station, test console).

- Mechanical Man - Physical opening and closing of on-board switches by mechanical carry-on equipment controlled by the checkout computer.

- Modification of Test Philosophy - This can take several forms, one of which might be performance of major portions of testing automatically without on-board switching and monitoring with a separate test to verify control and display performance and calibration.

5. The Astronaut/Computer Interface

Some thought has been devoted to the relationship between the astronaut and the ACE computer and to a progression of astronaut/computer interface design concepts. It was assumed that computer control of checkout would increase from the present Apollo concept to a "fully" automated concept. These improvements might take the following form in a time-ordered sequence:

a. Provide a piece of "carry-on" equipment that would allow an astronaut/spacecraft observer to insert a "yes" or "no" response directly into the computer. He could thereby affirm the receipt of instructions and the completion of activity sequences and could respond to questions concerning equipment status.

b. Provide an on-board, computer-controlled, program and instruction display such as a CRT, computer-stepped slide projector, movie projector, or TV. This would replace the present system of scripts, checklists, and verbal communications to provide the crew with test operational instructions, test step information, and test status. A system such as this might even be incorporated into the normal in-flight spacecraft inventory to be used for mission plans and operational checklists.

c. The final step in this progression would be to provide a keyboard computer input device to the above system to allow complete input/output communication between on-board crew and the checkout computer.

An interesting conclusion emerges from the preceding line of thought, namely, that in mission simulation testing and the final portion of countdown there is no requirement for an interface between the astronaut and ground computer beyond that which is already built into the spacecraft. This is true because the checkout computer, in this case, is acting as a slave to the astronaut through the normal spacecraft systems. A direct interface between astronaut and ground computer is required only in combined system testing.

C. Plans for the Second Quarter

The conclusions and hypotheses at which we have arrived as a result of the work accomplished to date are based, necessarily, on analyses of preliminary and limited data and upon initial discussions with MSC personnel. The main objectives for the second quarter are to collect and analyze additional and more comprehensive data on:

- The requirements for astronaut participation in performance of spacecraft missions.

- The requirements for astronaut participation in performance of prelaunch system checkout.

- Prelaunch checkout procedures and equipment design.

- The designs of the Apollo spacecraft crew stations and of the proposed interfaces between the astronauts and the prelaunch checkout system.

- Value of participation in prelaunch checkout as an astronaut training aid.

- The effect of on-board operator participation on test run time, failure diagnosis, and emergency procedures.

The following activities are planned in pursuant of these objectives:

1. Review Gemini and Apollo Documentation

The studies completed thus far have been based entirely on Mercury and very preliminary Apollo documentation. It now becomes necessary to verify the results of these studies with more recent Gemini and Apollo information. It is also important to delve more deeply into the crew tasks involved in specific operations of individual systems and of specific missions to modify and correct our concept of the crews' function, especially in the light of multiple-man crews. The type of information already requested from NASA includes Gemini and Apollo documentation in the three major areas of crew tasks, crew station descriptions or specifications, and test descriptions and procedures.

2. Astronaut Interviews

Having verified our concept of crew operational tasks through study of Gemini and Apollo documentation, we will obtain the astronauts' views on: (1) how they best learn these tasks; (2) in what types of tests they feel it is most useful to participate, from a training standpoint; (3) what types of system tasks they learn best from participating in spacecraft testing; (4) how else, other than by participating in spacecraft testing, they feel they could learn these tasks; (5) what types of information, other than operational information, they gain from participating in spacecraft testing; (6) what types of test participation, other than physically occupying the crew station, they

feel are important to their training and why; (7) how much cross-training multiple crews should receive; (8) how and when they know they are subjectively competent in their job performance; (9) how much participation in testing adds to their confidence in the spacecraft, the launch vehicle, and the launch crew.

The "user's" answers to these types of questions may identify particular tests in which it might not be necessary for them to participate or they may identify other means of training them in specific areas without their participating fully in testing. This idea of astronaut interviewing has been discussed with MSC personnel responsible for the planning of astronaut training. They were highly enthusiastic and promised their cooperation both in the review of interview designs and in the conduct of interviews. This interest appears to stem from some concern over the optimum amount of knowledge gained through involvement, both in and out of the spacecraft, during testing operations.

3. Apollo Design and Test Personnel Interviews

In order to balance the picture, it is desirable to conduct a similar interview for Apollo design and test personnel. Although design and test personnel are considered jointly, we realize that the interviews will be substantially different. We expect that the design personnel will be questioned as to: (1) how they would expect the astronauts to operate their individual systems; (2) how they feel the astronauts can best learn these procedures; (3) of what other information, besides operational procedures, the astronauts should be aware; (4) how best the crew can assist in failure diagnosis both during testing and in actual flight; (5) what knowledge regarding individual systems the crew should attempt to derive by participating in spacecraft testing; (6) ideas on prime system design to facilitate automatic testing with and without crew participation.

The questions to be asked of test personnel would be somewhat similar, but from a definite testing viewpoint rather than from a prime equipment standpoint. Regarding the questions on system design, for instance, they would be asked for ideas on test equipment design for automatic testing with and without crew participation. They would also be asked to comment on such things as the interface designs discussed earlier. These ideas would then be filtered and studied in further detail to determine their feasibility within state-of-the-art expectations.

We anticipate that these interviews can include questions which will provide data relevant to the concerns of Task I of this project.

Several additional studies will also be initiated:

- To identify major differences in system and mission operations, an examination will be made of the astronaut tasks involving the LEM as compared to the command module. This study will also examine the implications of the prelaunch accessibility of the LEM for astronaut participation in prelaunch checkout.

- A second study will seek to obtain an estimate of the gains and losses in test and setup time with various amounts of astronaut participation and with various astronaut/computer interface concepts.

- Another study will look at proposed countdown and checkout procedures to optimize computer usage with the crew aboard. It appears that procedures such as switch position checks might be computer aided to reduce time. For example, the astronaut could set switches by SOP with the computer monitoring these positions by program. If a switch was not set properly at a certain time, the computer would stop the count or flag the discrepancy at an STE console. The STE could then instruct the astronaut as to what switch was in error. This type of procedure would eliminate the calling out of each switch position over voice communication channels and having the crew affirm each. Some of these checklists become quite lengthy in a countdown.

TASK III. MAN-COMPUTER ROLES IN PRELAUNCH FAILURE DIAGNOSIS

A. Introduction

To date, prelaunch fault diagnosis has been primarily a manual and verbal process. The primary emphasis has been on human memory and the test engineers' ability to draw logical inferences between the occurrence of failure syndromes and the possible locus of a failure. For Apollo, however, it appears feasible to assign specific roles to the spacecraft crew members, the test conductor, and the system test engineers, and to the computer in such a way as to facilitate fault diagnosis. In addition, it should be possible to design certain types of job aids (perhaps controlled or mediated by the computer) which will help the personnel in diagnostic activities. Task III is a study of these roles and the resultant man-computer interface requirements.

The initial phase of this task was devoted to developing a definitive description of the context within which these diagnostic operations will be carried out, namely, the checkout operation as a whole. As a result, some of this work is also pertinent to the Task I study of the establishment of confidence as an objective for checkout operations. These initial studies have tended to emphasize the man-computer role in arriving at an acceptance decision, rather than in fault diagnosis per se, since this is the essential first step and since the rationale applicable to this decision is readily extendable to the more complex diagnosis decisions.

More specifically, the initial phase of Task III can be divided into work on the following subtasks:

1. Establishing the general context for the study, especially in relationship to the concepts being studied in Task I.

2. Clarifying and expanding the initial description of the study task.

3. Developing concepts and procedures which will guide the future work of the study.

B. General Context

The central problem in Task III is the determination of criteria for selecting diagnostic operations to be implemented by men, by automated equipment,

checkout (and this can later be extended to diagnosis) is a planned procedure. The operations, the sequence in which they will be performed, and the conditions under which various alternatives will be selected are all specified beforehand. In this process, man performs three essential functions:

- He develops the plan.
- He implements the prescribed operations.
- He introduces deviations from the planned procedure (or modifies the plan) in the course of implementing it.

The third role can be described more succinctly by stating that man introduces ad hoc deviations from the planned procedure.

The preceding discussion makes it clear that examination of the first of these roles is an essential aspect of Task I. The second role is a central consideration in Task III, and the third is a matter of mutual interest. The approach adopted for Task III, therefore, is to start with an examination of the way in which planned procedures can be implemented by various combinations of men and automated equipment. The first step is the development of a suitable description of the nature of the planned procedures.

C. Structure of Elemental Checkout Operations

The concepts and definitions used in decision theory provide a descriptive structure of the kind required, even though the algorithms of decision theory cannot be implemented when applied to problems of the magnitude and variability encountered in practical checkout operations. In the following discussion, therefore, it is understood that the reason for describing these procedures in decision theoretic terms is to obtain an orderly, comprehensible description of the operations of interest which will serve as a guide in future observation and analysis; it is not anticipated that the detailed analytical procedures of decision theory will be directly applicable to checkout operations.

The elemental operation in any checkout procedure is a sequence of four steps:

1. A test is carried out.
2. The results of the test are observed.
3. The results are interpreted.
4. The next action is initiated.

For general applicability, the word "test" in this description must be interpreted broadly: it is a collection of operations which can be called individual "tests" in normal engineering parlance. The most frequently used test operation consists of feeding inputs into a piece of equipment in order to observe its response. However, the test operation can also include direct visual inspection, subjection of the equipment to environmental stresses, detection of indirect "responses" such as operating temperatures, noises or vibrations, etc.

On a purely logical basis, Step 2, the observation of results, can be considered an integral part of Step 1, carrying out the test. It is specified separately because it provides a convenient way of describing the process more precisely. In a planned procedure, the possible results of any test can be specified beforehand. Thus all of the factors of interest in a formal analysis can be represented by defining a set of observations, Z, such that each $z \epsilon Z$* is one of the specified possible observations resulting from implementation of the test procedure. It can be assumed without loss of generality that the set Z is finite. Even if a particular observation (e.g., a voltage) may range over a continuum, the description of the output can be quantized by representing the continuum as a sequence of discrete intervals, and any desired accuracy can be obtained by making the intervals sufficiently small. In conformance with the preceding definition of "test," an observation, z, does not necessarily correspond with a simple "reading" but may be some combination of possible outputs. For example, it may consist of a particular set of signal light outputs; a particular combination of peak voltage, rise time, and decay time; the magnitudes of all the Fourier coefficients of a waveform; etc. One of the important possible variations from a planned procedure consists of making an observation not specified originally as a member of Z.

Skipping Step 3 for the moment, the test concludes with an action, Step 4. Some typical actions consist of:

(a) acceptance of the equipment as ready for use.

(b) diagnosis or replacement.

(c) additional tests on the same equipment.

---

*This notation is a shorthand indication that a particular observation is designated by z and that the collection of all possible observations is the set designated by Z: it is usually read "z in Z" or "z which is a member of the set Z."

The reason for postponing discussion of Step 3 is now apparent: each of these actions is predicated upon a particular interpretation of the observations. Action (a) is selected when the observation is interpreted as indicating that the equipment will operate satisfactorily. The need for Action (b) is indicated when there is some observation indicating a malfunction which must be located and eliminated. When the interpretation of the observation is ambiguous further investigation, as in Action (c), is undertaken. In other words, in a planned procedure, there is a specified set of actions, A, and a specific $a \epsilon A$ is designated for each possible observation (or for each $z \epsilon Z$); that is, a planned procedure includes the specification of a decision function, $d(z) = a$, which associates a specific $z \epsilon Z$ with each $a \epsilon A$ (i.e., the decision function maps Z into A).

As was illustrated previously, it is possible that there might be two different situations in which the same z is obtained but in which different actions are considered most appropriate. This is described formally by defining a class of decision functions, D, such that each $d \epsilon D$ maps Z into A. The distinction between the two situations just mentioned can be described succinctly by stating that a different decision rule, or way of selecting a $d \epsilon D$, is used in each of the two instances. Decision theory algorithms are procedures for selecting a $d \epsilon D$ which can be considered optimal when various specified kinds of information are available. The inapplicability of decision theory to practical checkout problems is due, in large measure, to the lack of availability of this information. However, this does not preclude the use of the concepts already defined in the descriptive manner being developed in this discussion.

D. Complex Checkout Operations

The "elemental checkout operation" can be represented diagrammatically as shown in Figure 1, which shows the two sets Z and A but not the decision function which links them. Practical checkout procedures can be regarded as extensive, complex combinations of these elements, and the character of the total operation depends critically on the nature of these relationships.

An example of such a combination of elements is shown in Figure 2. In this case, Z_1 can be considered the information set discussed in the preceding section. Then a_1 is the action which consists of initiating and carrying out the next step in the procedure. This action is selected when the $z \epsilon Z_1$ is in the subset interpreted as indicating that the test result is "satisfactory" and the result of this action is the new information set, Z_2, obtained from the next test in the sequence.

For other possible results in Z_1, the indicated action is diagnosis and repair, a_2. When this action is completed, more information is available as indicated by the set $Z_{1,2}$. The particular $z \epsilon Z_{1,2}$ obtained may indicate that the equipment is now satisfactory so the appropriate action is now a_1. If the difficulty is not

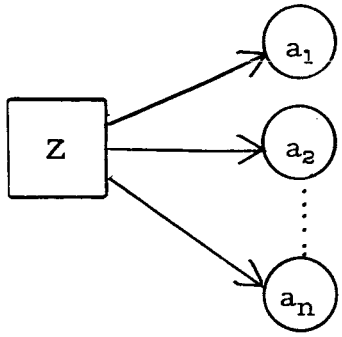


Figure 1. The Elemental Checkout Operation

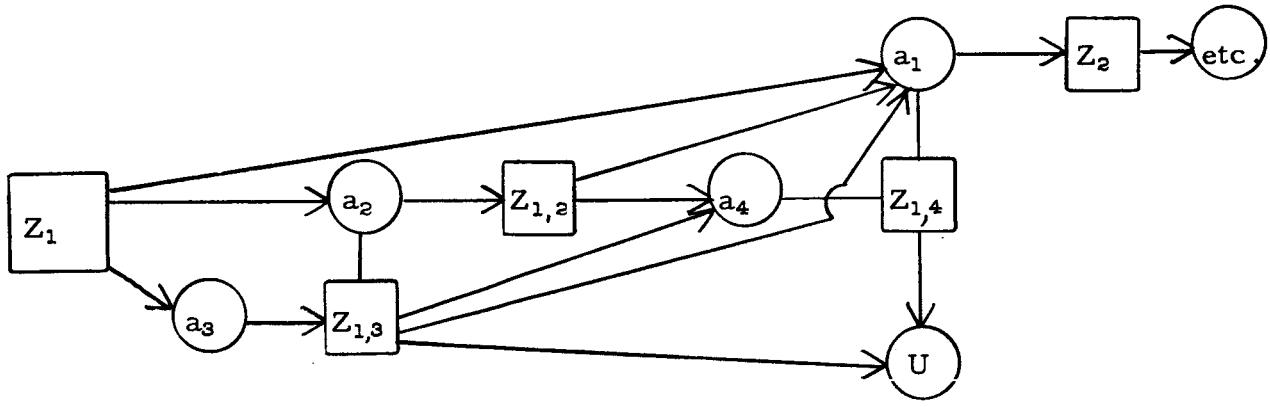


Figure 2. More Complex Checkout Operation

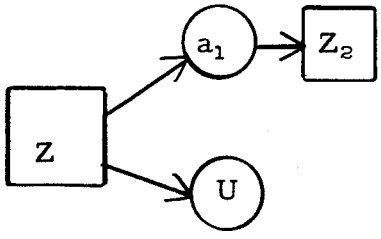


Figure 3. Alternative Representation of Operation in Figure 2

located, additional diagnosis and repair becomes the appropriate action as indicated by a_4 and its information set, $Z_{1,4}$. Again, the particular result achieved may lead to a_1. On the other hand, the equipment may still be malfunctioning. In this case, some additional action is required, but it may not have been planned, either because of lack of time, or because this particular contingency was not foreseen. The result is then the initiation of some unplanned action labeled U.

The final action, a_3, consists of performing additional tests different from those undertaken as diagnosis in a_2. The result may lead to acceptance, a_1, to some diagnostic and replacement action such as a_2 or a_4, or even to unplanned action, U.

The definition of an elemental operation provides for analytical flexibility since the designation of a particular portion of a complex sequence as elemental is arbitrary as long as the results of the actions involved can be represented as the production of new information sets. Thus the entire operation represented in Figure 2 can be represented as shown in Figure 3 by considering it to be an "elemental operation" producing the information set Z which is the union of Z_1, $Z_{1,2}$, $Z_{1,3}$, and $Z_{1,4}$. The representation selected in any instance can be tailored to analytical convenience and to the "natural" subdivisions of the physical procedure being considered.

E. Deviations from Planned Procedures

We have previously indicated that one of the human functions in checkout operations is the introduction of deviations from planned proceudres on an ad hoc basis (see paragraph B). These activities provide much of the basic data for study and analysis of the more complex human behavior and man-machine relationships in checkout and diagnosis and therefore we can discuss these in greater detail. In accordance with the discussion set forth in the preceding section, it is possible to list the following subcategories of ad hoc actions taken by man in checkout systems:

a. He may select a $d \epsilon D$ other than that planned.

b. He may augment A, adjusting the selection of $d \epsilon D$ accordingly.

c. He may augment Z, adjusting selection of d and structure of A accordingly.

d. He may alter the sequence in which elemental operations are performed, adjusting interpretation and action accordingly.

e. He may implement an unplanned procedure.

These are discussed in detail in the following sections.

1. Changes in the Decision Rule

The decision rule, d(z), specified in the original plan associates an action with each possible test result, z. However, when the test is run and the result observed, the man is still free to select a different action. In terms of the formal structure described above, he changes the decision rule.

There may be many rational reasons for introducing an ad hoc change of this type. As one example, a particular result may call for a_3, (additional tests) as described for Figure 2. It may happen, however, that the choice is made at a time when substituting a_2 (more detailed diagnosis) for a_3 will not delay the over-all procedure as originally expected (e.g., during an existing hold called for other reasons). The man might then decide to use a_2 instead because he realizes it is a more thorough test and he has an unanticipated opportunity to use it. This is typical since time, or the lack of it, is a most common reason for introducing ad hoc procedures.

A second possible basis for changing the decision rule is the realization, acquired after planning was completed, that the original interpretation of certain test results was erroneous. For example, the man may realize that a particular result, designed to indicate a "go" condition may be obtainable even though a certain portion of the system is not functioning. Accordingly, he might select an action like a_3 (supplemental testing) rather than the prescribed a_1 (accept).

A third possible basis for introducing an ad hoc change of this kind results from the fact that selection of the most appropriate decision rule frequently requires a consideration of risk. Thus if the mission objectives are changed after the checkout planning is completed an indication of "go" previously considered satisfactory may now be interpreted as only partially valid requiring some action other than a_1 as originally planned.

Changes in d are, in a sense, the simplest since they do not affect the structure of Z, A, or the operational sequence. Intuitively, it is anticipated that the study will show that these are the ad hoc changes which occur most frequently.

2. Augmenting the Action Set, A

This second category of ad hoc deviations from original plans occurs when the man introduces consideration of an action not described in the original set, A. This may happen as a result of invention. In the operating circumstances it becomes apparent that an action not previously considered is available,

feasible, and perhaps desirable. Unforeseen circumstances may suggest alternate actions. For instance, a spare part, not normally on hand, may just happen to be available and using it may be an action preferable to the planned diagnosis and repair. The introduction of a jury rig is an action in this classification. A closely related action is replanning a part of the mission to eliminate the need for a particular item of equipment.

Once the action set is augmented, the original decision rule is no longer adequate since it was selected without consideration of the desirability of the newly added action. Thus expanding A on an ad hoc basis requires an appropriate reconsideration of the original d.

3. Augmenting the Observation Set, Z

A frequently cited reason for favoring human participation in the detailed implementation of the checkout procedure is that he might notice some unanticipated but significant aspect of the equipment behavior. As one instance, a voltage reading might be part of the observation set, but the human might notice a momentary fluctuation occurring before the voltage is achieved. He will then have made an observation not specified in the set Z. There are also many simple observations in this class such as hearing a strange noise, feeling the heat from an overheating part, sensing unusual resistance in manipulating a mechanical part, or just seeing a dangling wire. Each such observation is not usually a member of the planned set of observations, Z.

Intelligent utilization of these unexpected observations requires, first, a modification of the decision rule since the observed z is not in its range of definition, and possible modification of the action set as well since some unanticipated action may be required to diagnose and remedy the condition leading to the observation.

4. Changes in the Operational Sequence

One of the principal advantages of maintaining active human control over the checkout is the resultant flexibility with which the various operational subroutines can be executed. In particular, when one subsequence becomes overly time-consuming because it requires a difficult diagnosis or repair operation, the over-all time lost can be minimized by carrying out other tests simultaneously.

Changes of this time can become very complex. Formally, each new observation set can include all of the observations which preceded it. Thus changing the order in which the elemental tests are performed can also change the character of the elemental tests themselves. In practice, of course, there are many instances in which this is not the case: the information obtained from

preceding tests may have no significant influence on the selection of an action in the subsequent test. In these cases the two tests can be called "independent," and the order in which such independent tests is performed is clearly immaterial.

Thus, there are two subclasses in this category. The first is changing the order in which independent tests are performed, which can be done arbitrarily. The second is changing the order in which there is some interdependence between tests: this creates a need for some replanning of the types already mentioned.

5. Implementation of Unplanned Procedures

The final category of ad hoc human activity results from the fact that practical checkout procedures are never completely planned. Within a reasonable time and with reasonable effort, only a limited number of contingencies can be foreseen and accounted for. The events least likely to occur may be overlooked, or even deliberately ignored. A typical instance is shown in Figure 2 where a_2 and a_4 presumedly exhaust the planned diagnostic action, but with the possibility left open that a detected malfunction will still not be corrected requiring additional but unspecified action. In these cases, the man works out procedures as he goes along, attempting to return to the planned sequence as expeditiously as possible.

6. Evaluation of Ad Hoc Procedures

The structure described above provides a basis for investigating the nature of ad hoc procedures, the frequency with which they occur, and finally, the consequences of using such procedures. These consequences are of two kinds, immediate and long term.

The immediate consequences are those affecting the operation which was in process when the ad hoc deviation was introduced. A qualitative evaluation can be obtained by determining the answers to such questions as: did the introduction of the ad hoc deviation accomplish something which could not have been done by merely following the planned procedure? If so, what descernible effect did this have on the success of the over-all operation? What might have happened if the procedure had not been used? Did introduction of the ad hoc procedure increase or decrease the operating time significantly? As case studies of this kind accumulate, it will gradually become possible to determine if the immediate consequences of various categories of ad hoc activity are desirable, undesirable, or insignificant.

The immediate consequences are not the only pertinent consideration. It is also necessary to investigate the effect of ad hoc deviations on subsequent operations. The only valid reason for the introduction of such a procedure is

compensation for a deficiency in the original plan. In some instances, this deficiency may be significant only in the particular combination of circumstances in which the incident occurred. In other cases the deficiency may be generally significant, and the new procedure, or some more carefully studied variation, may be incorporated in the planned procedures for subsequent operations. In this way, the ad hoc experiences become an important element in the group learning process which guides the evolutionary development of improved checkout procedures. In fact, the incorporation of an ad hoc procedure in later formal plans is the best evidence for its desirability.

F. Plans for Next Quarter

The concepts described above have been used to organize and plan the work which will be undertaken in the next phase of Task III. This will consist of detailed investigations of past and present checkout procedures with the following specific objectives.

1. The initial considerations which, for purposes of clarity, were limited to the checkout problem (in which only one hypothesis is evaluated), will be extended to include diagnosis, which is characterized by the existance of multiple hypotheses (or action spaces with more than two alternative actions).

2. Specific examples of checkout procedures will be selected and described formally in order to confirm the utility of the suggested structure and establish a factual basis for further analysis.

3. These examples will be analyzed to determine the nature, frequency and significance of the ad hoc deviations categorized on preceding pages.

4. These examples will also be analyzed to determine the nature of the data processing operations (man or computer) required to implement the planned procedure, with these operations categorized as:

   - those required to collect information,
   - those required to disseminate information,
   - those involving the use of standard algorithms for numerical or logical manipulations,
   - those involving logical decision making,
   - those involving information retrieval.

TASK IV. FAILURE ANTICIPATION

A. Introduction

On this task we have been investigating possible techniques of failure anticipation. The fundamental problem consists of two major parts: (1) identification of test responses and data transforms (if any) which can assist in predicting anticipated incipient failures, and (2) development of efficient techniques for developing and displaying this knowledge to the test engineer for subsequent action.

Manual checkout systems usually provide the test engineer with complete indications, including graphic recordings, of all test responses. Given this information and the relatively slow pace of manual operations he is able to make inferences based on "minor" clues (e.g., flickering meters, slight drifts within limits, etc.) which may be indicative of failures at some future point in time. But even here, the test engineer is limited by incomplete records, by an inability to perform complex calculations rapidly, and by the fact that he must work with a limited number of test responses (limited in terms of the number of points sampled, the number of repetitions of individual measures, and the range over which any parameter is measured).

As we move towards more automated checkout systems, we gain the advantage of large memory capacities which permit collection and temporary storage of more complete test records and the ability to perform complex calculations in short order. This would permit us to analyze response data for prediction purposes if only we knew which data transforms to select. Since it is not feasible to evaluate all possible parameters or combinations for even a single system, it is first necessary to identify transfer functions and parametric analyses which have predictive value and the systems and circumstances to which these are applicable. This can be established via the following sequence of activities:

1. Development of conceptual and mathematical models.

2. Use of these models on existing test data to evaluate their applicability and limitation.

3. Validation of the model's predictive value.

4. Determination of relative roles of man and computer.

5. Refinement of these techniques and integration into existing computerized checkout procedures and equipment.

These then represent the basic activities included as part of this task. (Steps 3, 4, and 5, will be undertaken only if the earlier steps result in meaningful progress.) The probability of developing reliable failure predictors is relatively low, in view of the singular lack of success in previous investigations of the problem. However, we feel that the high potential pay-off warrants continued efforts along this line which may unearth new and fruitful techniques or approaches. The current effort is being structured as an exploratory one and the effort devoted to it will depend on the degree of success we experience in its early phases.

The topic under investigation bears little resemblance to the more frequently performed studies on various phases of reliability estimation. Hence, it was first necessary to define the scope of the research and to evaluate the possible approaches which might initially be employed. While we are still in the early stages, it is now possible to discuss certain areas that show promise and should be further explored and others that appear to hold little prospect of success. The remaining portion of this chapter will present descriptions of our general thinking on the problem, the directions in which we have probed, and the opinions and conclusions which mark the progress to date.

B. Scope of The Problem

Generally speaking, this study will focus on tests conducted after delivery to AMR; the object being to determine if failure can be anticipated prior to lift-off.

This is quite different from and must not be confused with a statement to the effect that the probability of an item failing in some specified period has some numerical value when seemingly "all's well." Such risks fall within the domain of reliability estimations involving inherent failure rates or appropriately modified failure rates and are wholly unrelated to the subject under discussion. (This is not to say that philosophies on failure are agreed in this distinction; in fact, strong controversies have raged and color the entire issue. But this is not the place to attempt resolution of such issues.)

While the information generated within the period between AMR acceptance and lift-off carries the greatest weight in failure anticipation in the present context, it seems logical that information available prior to that period should also contribute to the system knowledge wherever such be feasible and seemingly worthwhile. This additional information will mostly serve the purpose of cross-referencing, checking historical data and verifying any anticipated failure.

There are certain kinds of failures that do not lend themselves to incipient failure analyses, and, in general, such areas pose unique problems which will not be considered, at least in the early phases of this task. These include:

- Parts and components (e.g., transistors, capacitors) that simply fail between one test and the next will elude even the most sophisticated failure anticipation techniques. These are still characterized as random failures until such time as an underlying cause can be assigned to them, and therefore are best handled by standard reliability techniques.

- Human errors that lead to failures are likewise unpredictable in the present sense. It seems unlikely that every function performed by the human operator can be mechanized or cross-checked and verified for correct operation and as yet we still do not have any precise way of assessing human reliability (i.e., predicting the nature and frequency of his errors).

C. Approach

In order to establish a starting point and a plan for subsequent activities a preliminary survey was undertaken to determine what information is currently available on space vehicles that are in various stages of completion. The information we sought may be itemized as follows:

- Data on prelaunch checkout operations for previous space flights and booster launches.

- Drawings, schematics, block-diagrams, flow-diagrams which show system complexity and interconnections.

- Failure reports and failure modes and effects analyses.

- Other existing data describing acceptance, environmental and integrity tests.

- Data on the nature and accuracy of test measurements.

- Post-lift-off and in-flight telemetry information.

To date we have received descriptive information on certain spacecraft and booster systems and the associated test procedures, but no test result data. The absence of such test data has precluded any attempt at tracing the system test history pertaining to actual incidences of failure. Our

approach during this quarter has therefore been restricted to the development of conceptual and mathematical models which we hope to test as soon as test data become available.

Descriptions of the statistical techniques that were analyzed as being particularly appropriate to information collected thus far are presented below. The principal results are included to the extent of their development during this quarter.

D. Specific Statistical Techniques

1. Cu-Sum Technique

Many parameters are measured to provide information about performance of a system. This information then results in a decision either that the system is essentially stable or that some change has occurred. A useful statistical analysis in relation to this problem is the application of cumulative sum methods.

The cumulative sum analysis provides a means of presenting, in time, the apparent mean of any group of consecutive data points, e.g., from a continually monitored system or from a periodically tested system. The cumulative sum chart consists of observing the cumulative sum

$$S_N = \sum_{i=1}^{N} x_i$$

against N, the total number of samples recorded and x_i, the individual sample deviations from a target value. Since the mean, relative to a target value T, of the last r points is $(S_N - S_{N-r})/r$, the mean of any group of consecutive points is proportional to the slope of the line joining the first and last points of the group. A change in mean value of the original parameter will be indicated by a change in slope on the cumulative sum representation, including the number of results relevant to the determination of this change.

For the cumulative sum analysis, if we wish to detect changes in one direction (a trend) then decision criteria can easily be established. Unlike the usual process control schemes,with arithmetic or geometric (exponentially weighted) running means, the cumulative sum analysis is quite sensitive to moderate changes in the mean value, i.e., moderate changes in mean value will be readily reflected in changes in the slope. In addition, decision criteria can be calculated thus: If the mean value corresponding to zero slope is called the reference value k, which may or may not be the target value T, and we have a decision interval h, then a decision that a change has occurred is taken if

$$\sum_{i = N - r + 1}^{N} (x_i - k) \geq h$$

where $r \leq N$.

Moderate shifts may be roughly understood to be in the range of 0.5σ - 2.0σ. A shift of this magnitude is most efficiently detected by this technique as opposed to the standard Shewhart-type control analyses. Décisions based on indications of shifts less than 0.5σ may lead to unnecessarily frequent and false actions. Accordingly, the average length of detecting such a shift is longer for this technique than for standard Shewhart control analyses.

It may also be emphasized that the point at which a shift occurred is vividly noticeable in the utilization of this technique, permitting the precise determination of the time at which changes indicative of failure anticipation occur.

By utilizing the analytical technique of the cumulative sum analysis and its associated decision criteria a powerful trend analysis tool can be brought to focus on the problem of failure anticipation.

2. Peak Hazard

The applicability of peak hazard techniques in relation to the analysis of "failure bunching" is being investigated. The concept being considered is that there exist relatively long periods of failure-free functioning followed by a short "failure period" during which a temporarily severe environmental condition will weaken or destroy components, and therefore cause system failures. Following a "failure period" there will be another period during which no failures occur. The environmental condition of the failure period may not cause truly simultaneous failures to occur, but a bunching effect may be observed.

To illustrate, let v be the normal independent failure rate of components and λ be the rate at which environmental shocks occur. Let p be the probability that an environmental shock is effective in causing component failure and N be the number of similar components in the system. Then one formulation of the expected number of failures due to environmental shocks in time t is

$$\xi t = Nt[v + \lambda p]$$

when ξ is the estimated expectation of system failure rate. The probability of system survival, based on an exponential model, in time t is

$$e^{-\xi t} = \exp\left[-t\left(Nv + \lambda\left[1 - (1-p)^N\right]\right)\right]$$

where $\lambda\left[1 - (1-p)^N\right]$ is the rate at which effective shocks occur and $\left[1 - (1-p)^N\right]$ is the probability that such a shock causes failure of at least one of N components. By letting $\lambda = 1$ we can restrict the model to one particular environmental shock in a time interval.

This model is only approximate and illustrative and assumes the failure period to be of vanishing duration. Additional bunching of failures may occur when one component induces failure of others or components "wear out" together.

One direction in which models of failure bunching may be generalized is the following. Assume that the occurrence of an environmental shock renders the C active components (the number required for the system to function) simultaneously susceptible to failure, and the probability that a component fails, given a particular type of shock, is p, independent of previous history. If shocks occur at independently distributed time intervals with d.f. F(t), then the probability that the time to failure (T_s) of s or fewer spare components will exceed t is,

$$\Pr\left[T_s > t\right] = \sum_{n=0}^{\infty} \left[F^{n*}(t) - F^{(n+1)*}(t)\right] \sum_{j=0}^{s} p_j^{n*}$$

where $F^{n*}(t)$ denotes the n-fold convolution of the d.f. of the inter-shock time intervals, and p_j^{n*} is the n-fold convolution of the d.f. of the number of simultaneous failed components when shock occurs.

As an illustration, if shocks occur according to a Poisson process with rate u, we have

$$\Pr\left[T_s > t\right] = \sum_{n=0}^{\infty} e^{-ut} \frac{(ut)^n}{n!} \sum_{j=0}^{s} p_j^{n*}$$

and

$$E(T_s) = \frac{1}{u} \sum_{n=0}^{\infty} \sum_{j=0}^{s} p_j^{n*}$$

The hypothesis that environment, and hence failure rate changes rapidly and independently may not always be reasonable. For example, different periods of hazard may be associated with period when equipment is operating, and when it is quiescent.

The above ideas may be related to various phases of integrated systems checkout testing, and therefore, seem useful in the present context of failure anticipation.

The present plan is to continue further exploration of this process as associated with testing duration and on/off time of equipment in relation to the notion of "failure periods."

3. Correlation and Regression Techniques

The applicability of multiple regression techniques, utilizing autocorrelation and cross-correlation functions, is being investigated in relation to failure anticipation of spacecraft systems. Those systems which have either two or more identical components or subsystems which perform the same function or redundant subsystems are being studied. The hypothesis to be tested is that n components of identical design and manufacture which are to perform and function in exactly the same manner should exhibit, under identical testing, equal test outputs (results) within prescribed accuracy limits. Therefore, the cross-correlation coefficients of test results should be approximately one and the autocorrelation distributions should be comparatively homogeneous.

In the testing of each individual component or subsystem, the output data may be considered as N + k observations from a stationary process as a function of time. We can denote the observations as $x_1, x_2, \ldots x_{N+k}$ where x_{N+k} is the last observation. These observations could be N values from an analog record or discrete points in a digital process. The autocorrelation analysis is performed by computing the same autocorrelation function at k points ($k < N$) defined by

$$R_{xx}(i) = \frac{1}{N} \sum_{j=1}^{N} x_j x_{j+i}, \quad i = 0, 1, \ldots, k.$$

A more convenient form for application is the normalized autocorrelation function. This is defined by dividing $R_{xx}(i)$ by $R_{xx}(\ )$, giving

$$\Gamma_{xx}(i) = \frac{R_{xx}(i)}{R_{xx}(0)}, \quad i = 0, 1, \ldots k$$

where $\Gamma_{xx}(i)$ are the correlation coefficients bounded as follows

$$-1 \leqq \Gamma_{xx}(i) \leqq 1, \quad i = 0, 1, \ldots k.$$

If extrapolation of the time series is desired, then a regression analysis can

be utilized. Therefore, to predict one time period (t + 1) in the future, we would solve the equation

$$\hat{x}_{t+1} = b_1 x_t + b_2 x_{t-1} + \cdots + b_k x_{t-k+1}$$

where the b's are the regression coefficients to be computed.

The above technique can be extended to n identical components or subsystems by utilizing the cross-correlation functions of the output data. For the case of two variables, x and y, the technique may be used to predict one variable as a function of both observations on itself and observations of the other variable. The cross-correlation coefficient of x with y is given by

$$\Gamma_{xy}(i) = \frac{R_{xy}(i)}{\sqrt{R_{xx}(0)\, R_{yy}(0)}}, \quad i = 0, 1, \ldots k$$

where $R_{xx}(0)$ and $R_{yy}(0)$ are the autocorrelation functions of x and y respectively, and

$$R_{xy}(i) = \frac{1}{N} \sum_{j=1}^{N} x_j\, y_{j+i}$$

It is obvious that $\Gamma_{yx}(i)$ can also be computed.

If extrapolation is desired, a regression analysis can be performed utilizing the regression equation

$$\hat{x}_o = b_1 x_1 + \cdots + b_r x_r + c_1 y_1 + \cdots + c_s y_s$$

where

r, s = at most N + k points.

The regression coefficients b_r and c_s are found by minimizing the appropriate sum of squares and solving the resulting set of simultaneous linear equations. The method of evaluating the regression coefficients is not indicated herein. This method can easily be extended for $n > 2$ variables, i.e., n identical components or subsystems.

Once the auto- and cross-correlations have been established, statistical tests of significance, utilizing any desired confidence interval, can be applied to determine if (1) significant differences between correlation coefficients are evident, (2) one or more of the n components is exhibiting unusual or

erratic performance, and (3) the cross-correlation coefficients are significantly different from 1.0. In addition, regression analysis can be used for extrapolation of performance. This extrapolation is one form of trend analysis. An example of these techniques is given below.

Example

A problem of interest is to evaluate value pressure drop data as a function of propellant pressure. For example, a pressure transducer may be mounted on the main propellant feed line to measure propellant pressure to three engines. In addition, the pressure drop for each engine valve may also be measured. An operating test is performed and all observations are found to be within tolerance limits. The obvious decision is that all systems are functioning properly.

However, valuable information as to the future state of the system may be lost if the inherent properties of the available data are not fully utilized. This is shown by the following example. Let w(t), x (t) and y(t) represent readings of valve pressure drop across the valves and z(t) represent propellant pressure. The normalized cross-correlation function of w(t) with z(t) might be illustrated as

$$\Gamma_{wz}(i) = \frac{R_{wz}(i)}{\sqrt{R_{ww}(o)\, R_{zz}(0)}} \quad , \; i = 0, \; 1, \; \ldots \; k$$

and in the same manner we can compute $\Gamma_{xz}(i)$ and $\Gamma_{yz}(i)$.

We can then test the hypothesis that $\Gamma_{wz}(i) = \Gamma_{xz}(i) = \Gamma_{yx}(i)$. Utilizing these statistics gives rise to, at least, three cases:

Case 1

The hypothesis that the sets of correlation coefficients $\Gamma_{wz}(i) = \Gamma_{xz}(i) = \Gamma_{yz}(i)$ is accepted as true, and in addition, we compute $\Gamma_{zz}(i)$; and, if it is sufficiently close to 1.0, we may accept our initial decision that all systems are functioning properly.

Case 2

The hypothesis that $\Gamma_{wt}(i) = \Gamma_{xz}(i) = \Gamma_{yx}(i)$ is not accepted as true, and again Γ_{zz} is sufficiently close to 1.0. We may then examine our data and find one of the three valves is erratic in operation in comparison to the other two. In this case, the possible decisions are:

a) Rerun test

b) Perform diagnostic procedure

c) Use regression analysis to anticipate a failure trend in time:

$$x_{t+k} = b_1 x_{t-k} + b_2 x_{t-1-k} + \cdots + b_n x_{t-n+k}$$

d) System is still functioning properly

Therefore, our basic decision has changed to the extent that one valve, while in tolerance, is not functioning completely adequately.

Case 3

The hypothesis that $\Gamma_{wz}(i) = \Gamma_{xz}(i) = \Gamma_{yz}(i)$ is accepted, but the values are all high or low in respect to the expected value of pressure drop readings. In this case, a decision to examine the propellant feed system may be forthcoming, since consistency of valve readings has been observed.

Conclusion

Other possibilities beyond the three cases presented exist, but these cases serve to illustrate that additional information pertaining to the problem of failure anticipation is often available. These data, if properly displayed and utilized by the test engineer, can lead to more meaningful prelaunch checkout of identical or redundant systems.

E. Plans for the Second Quarter

During the subsequent period we will be engaged in the following activities (subject to the limitations imposed by data availability):

1. Further development of the statistical techniques discussed previously with a view to applying them to actual systems which will undergo prelaunch checkout testing. This will consist of:

   . Examining the power and suitability of these techniques to prelaunch checkout data and systems generated on previous launch and space vehicles; and

   . Adapting these techniques to new and more general classes of information and data for advanced space vehicles.

2. Where failures have been observed in previous prelaunch checkout programs, these failures will be traced, to where checkout begins, to determine if any revised form of data collection, analysis, or monitoring would have successfully anticipated or averted such failure. It is quite conceivable that the effort will result in the formulation of additional statistical techniques.

3. The use of computers in the reduction and analysis of prelaunch checkout test information, within the framework of the statistical methods developed, will be studied. Attention will also be paid to the most practical and useful formats in which the results of the analyses are to be presented to test personnel.

4. The application of those of man's failure anticipating abilities which are not concretely definable in relation to prelaunch checkout testing will be investigated.

At the present time there exist wide gaps in the information and data required for satisfactory continuance of various phases of the project. Data collection in the following areas will be continued:

. Detailed prelaunch checkout tests and sequences on a time scale from arrival at AMR to countdown

. Detailed description of selected tests

- Measurements recorded and accuracy, by test
- Tolerance limits
- Test conditions
- Test steps

. A measurement list which can be cross-referenced to a particular test(s)

- Descriptions of methods of recording and storing of data
- Samples of detailed prelaunch checkout data on Apollo Boilerplate 13 launch
- Selected samples of detailed data on prelaunch checkout at Saturn launches
- Detailed description of tests and measurements of NAA combined system tests
- Subsystem failure modes and effects analyses for selected subsystems
- Samples of NASA failure reports on Apollo, Gemini, and Saturn launches.

TASK V. FUTURE MAN-MACHINE ROLES IN SEMIAUTOMATED TEST PLANNING AND PROGRAMMING

A. Introduction

Checkout computers appear to have great potential as tools for use in off-line test planning and programming and in dynamic on-line revision of the test plan to meet testing contingencies. In order to realize this potential, it is first necessary to explore the rules and constraints under which test plans and procedures are formulated and integrated into test sequences. Once this has been done the next step involves the development of techniques and guidelines which provide efficient two-way interaction between the test engineer and the computer. Both of these steps are being examined as part of this task.

We have identified some of the test planning activities in which computers can assist the test engineer. These include: (1) optimization of test sequences, scheduling of simultaneous tests, and sharing of support equipment or facilities; (2) development of test strategies and criteria; (3) on-line planning and control of test recycling and "back-off" in the event of holds or failure indications; (4) prediction of anticipated hold times for various interruptions in the test cycle; and (5) establishing requirements for a request language to permit discourse between test engineer and computer. Since some of these areas are also being investigated by The RAND Corporation, close coordination will be required to avoid duplication of effort.

First quarter activities on Task V have been limited to task organization and planning, review of available documents, and some preliminary analysis of test recycling, a problem which is of great interest to Apollo program personnel at both OMSF and MSC.

B. Task Organization

The activities relating to Task V have been organized into a logical sequence of subtasks as listed below. These constitute a plan for performance of Task V.

. Subtask 1. Analysis of Test Planning Activities

This subtask will include a study of current documentation pertinent to test planning such as "Operations Plans," "Test Plans," "Scripts," and other related material. It will also include discussion with personnel responsible for the creation of test plans and collection of data on current task planning activities. This investigation will explore and structure the concepts behind the relevant terms involved, e.g., test planning functions, planning rules, test

strategies, test methods, common test languages, on-line planning, adaptive checkout, design test programs, operational test programs, test efficiency, test depth, test programming flexibility, and interrelationships behind these concepts. The data will be organized and categorized in accordance with the various testing rules, constraints, and techniques identified during the conduct of this subtask.

. Subtask 2. Evaluation of Alternative Approaches to Man/Computer Roles in Off-Line and On-Line Test Planning

The second subtask will result in a more detailed examination of the concepts defined in Subtask 1 and an evaluation of the alternative approaches with special attention to efficiency, cost versus elapsed time, modification techniques (both on-line and off-line), modular construction for ease of expansion, programming problems, and training problems associated with each alternative. It will result in recommendations for those tasks which are best performed by a computer, and those test functions which are best retained in a manual, or mechanical, mode under the control of the design and test engineers. Both technical and philosophical justifications will be identified.

. Subtask 3. Development of Design, Training and Programming Requirements for Selected Concepts

The recommended allocations of functions, equipment capabilities, and operational procedures will be used to develop the detailed design, programming, and training requirements and schedules needed to implement these concepts.

. Subtask 4. Determination of Requirements of a Request Language to be Used by Test and Programming Personnel

This subtask will include a review and evaluation of request language requirements, as derived from test engineer needs. This will result in a set of requirements which can be compared with existing higher language type compilers and test languages to determine their applicability to prelaunch test planning and operations.

C. Review of Documents

During this reporting period, efforts on Subtask 1 have been concentrated on the procurement and review of documents which describe current plans for automated prelaunch checkout of Apollo and Saturn systems and the constraints relating to such operations.

The review of these documents has furnished the preliminary groundwork for exploring and defining some of the concepts behind automated test planning. As indicated in the performance plan for Subtask 1, these concepts will be developed and their implications for future operations discussed in the Task V Report. Presently available documentation on automatic checkout techniques deals primarily with the launch vehicle; therefore, it would be premature to draw any general conclusions before collecting more data. However, a need for more automatic techniques to facilitate test planning does appear to exist; first to assist the test designers/engineers at Marshall Space Flight Center in developing initial test requirements, and secondly to facilitate rapid, flexible test modification during on-line testing at the John F. Kennedy Space Center. Further investigation will be conducted in both areas.

D. Recycling of Test Steps

Recycling is the interruption of a pre-planned test routine in order to return to some earlier step in the sequence. It can be initiated (manually or automatically) for either of two reasons: (1) to re-trace a given test step which yielded ambiguous or suspicious data; or (2) to return the space vehicle and/or checkout system to a safe condition in the event of failures which create safety hazards or excessive stresses or which expend needed energy or equipment life.

Task V is focused primarily on the latter type of recycling since it establishes its own unique test planning requirements. (Recycling for the first reason is discussed in Tasks I and III which deal with confidence and fault diagnoses, respectively.)

The basic steps involved in establishing recycling requirements and procedures are:

1. Specification of the space vehicle and checkout system failures (or contingencies) which can occur during a given system test sequence.

2. Determination of the hazards (immediate or potential) inherent in each such failure and enumeration of the time available for compensatory action.

3. Identification of the point to which the test should be recycled and any remedial action which must be taken.

4. Estimation of the probable hold time or delay associated with such recycling (where possible) and enumeration of the other test activities that could safely continue during such periods.

5. Implementation of procedures (manual, automatic, or combinations) to control recycling and collateral testing.

To our knowledge these steps as yet have not been followed in developing any existing recycling guidelines. All of the test contingency plans reviewed to date appear to have been prepared specifically for a single test sequence on a single system (or combination of systems). While detailed recycling procedures obviously must be prepared on a test-by-test and system-by-system basis, it appears that there should be certain general guidelines applicable to classes of tests and systems.

If we consider the test paradigm as a branching sequence of activities (e.g., connection of a stimulus generator, initiation of a stimulus, etc.) followed by events (e.g., evaluation of a response) which determine the next subset of activities, it becomes apparent that recycling decisions can only be made at certain "choice points" during the sequence. The initial task in studying the recycling problem is the identification of these choice points.

Failure mode analyses of spacecraft are currently being prepared by the various Apollo system contractors, and it is to be hoped that these documents can be used in conjunction with existing test plans and programs to pinpoint possible classes of failures for each system and the choice points inherent in each test routine.

Once these points have been identified, it is necessary to determine the proper branch to be followed as a result of potentially negative status information available at that point. Possible choices include: proceeding to the next subset of activities in the sequence; holding at the present point and generating an alarm of some kind; retracing to some previous point in the sequence and then holding; shutting off all power, pressure or stimulation; activating some other portion of the checkout or support equipment; etc.

If recycling is indicated, it is next necessary to identify the test conditions to which the systems should be returned. This point can be determined on the basis of two major criteria -- elimination of hazards and minimization of hold time. The former is the more important, and thus must be considered first. For certain types of systems and tests (e.g., those involving fluids under pressure) it may be possible to back-off in stages with specified decision rules governing the amount of time at each stage. The second criterion, which becomes of interest once hazard and reliability considerations have been satisfied, may actually dictate recycling to a point still earlier in the sequence. This may be desirable to facilitate access to faulty equipment or diagnostic procedures.

The above analyses result in the development of contingency rules and procedures specific to a given system and test. The final steps involve the

integration of such procedures across systems and tests to optimize the efficiency of prelaunch checkout operations and establishment of techniques for implementation, control and monitoring of recycling.

E. Second Quarter Plans

Plans for the second quarter include continued review of documents and a trip of approximately one week's duration to the Marshall Space Flight Center at Huntsville, Alabama, and to the J. F. Kennedy Space Center at Cape Kennedy, Florida, to discuss test plan generation, implementation and revision. Discussion will include information requirements, personnel functions, sources of information, types of decisions involved in test planning, and problem areas as identified by test designers, test engineers, system design, programming, and computer engineers. This trip will also provide the opportunity to observe directly several of the stages of test planning and implementation. The completion of Subtask 1 has been scheduled for the second quarter.

During the next period we will also begin a more intensive investigation of the recycling aspects of test planning. Our plan is to select one spacecraft system for initial study in order to develop both specific recycling recommendations and more general guidelines applicable to other systems. This work will be done in close cooperation with the Apollo ACE group at MSC and OMSF.

SPECIAL TASK: INVESTIGATION OF APPARENT MOVEMENT PHENOMENA ON CRT DISPLAYS

During our study of display requirements for prelaunch checkout (conducted under subcontract to The RAND Corporation as part of NASA Contract NASr 21(08)), we observed the existence of certain apparent movement phenomena on CRT display units planned for use in Apollo prelaunch checkout. These phenomena were first observed during installation and experimental use of a dd51 unit at the J.F.Kennedy Space Center and later during factory testing of a dd60 and a dd74G unit.

These were first reported to NASA in February of 1964, and in June the Manned Spacecraft Center and the Office of Manned Space Flight, NASA, requested Dunlap and Associates, Inc. to study these phenomena and to conduct experiments to investigate their effects on operator performance. During the first quarter of the current contract, we have analyzed these visual effects in an effort to explain their underlying mechanisms and we have conducted a set of experiments to evaluate their impact on performance.

A. The Nature of the Phenomena

1. Flicker Fusion and Spatial Stability

It appears useful to separate general visual phenomena caused by intermittent light sources into two classes -- those occurring above and those occurring below CFF (Critical Fusion Frequency).

An intermittent light occurring below CFF is seen to flicker. The human visual apparatus is able to distinguish light and dark periods, and it interprets them (completely or mainly) in terms of temporal change. The apparent simplicity of the situation is belied, however, by the fact that the observer perceives a light to flicker at or about 20 cps, regardless of the objective flicker rate. Some variation is perceptible, as demonstrated by the finding that most observers can discriminate three subjective flicker rates: fast; medium; and slow. In general there is a concensus that a flickering display is to be avoided, and we may be sure that very little flicker occurs over 50 cps. Under the conditions of display which obtain on a typical fast-phosphor display, very few flicker effects remain over 40 cps; however, Dr. Bowen has noted some lingering flicker effects on the dd60 when the PRF rate was 50. These effects occurred in the periphery of vision, which has higher CFF rates than has the center of vision.

The CFF point is determined by a multiplicity of factors which combine in complex ways. Primary factors are brightness area, and duty cycle; and the total range of CFF values extends from about 3 cps to about 55 cps. In general the CFF point is driven higher with increases in brightness and in area illuminated and with smaller light/dark ratios. It happens that the characteristics which are desired on alphanumeric displays (e.g., on the dd74G and -H) are exactly those conditions which are conducive to high CFF values; hence it becomes difficult to eradicate all traces of flicker.

Until recently the literature led one to believe that lights flickering at above fusion frequencies were perceived simply as steady lights. The integrative mechanisms of the retinal-neural system were assumed to be sufficient to blend the photic pulses. However, as common observation of "moving" neon signs will demonstrate, the situation changes where there is discontinuity in both time and space. Apparent movement can occur when there is some intermittent illumination of separate elements.

It is disappointing to report that, with one exception, the literature on this topic is both extensive and vague. The one exception is a study by Crook and Wade* in which the visual instabilities in "refreshed" CRT displays were noted and an experiment performed to determine whether they degraded human performance. The results were negative.

It is evident that the perception of apparent movement is a complex phenomenon and depends almost as much upon the conditions of the observer as upon the conditions of observation. Some kind of movement may be perceived when two different parts of the retina are stimulated in succession -- in the limiting case by instantaneous succession. It appears that the human visual system is capable of differentiating spatio-temporal brightness discontinuities and that this forms the basis for the phenomena with which we are concerned.

2. Description of the Phenomena

Observation of a variety of display formats on the dd60 equipment (P-31 phosphor) led Dr. Bowen to categorize the phenomena into three classes.**

---

* Crook, M.W. and Wade, E.A. Effect of Periodic Luminance Reintensification on the Reading of Visual Displays. Tufts University, Medford, Mass., Aug 60.

**Since this writing a fourth has been reported, namely the "beat" effect due to the interaction of a single, bare, 60-cycle fluorescent light used for ambient illumination with the 50 cps refresh rate of the CRT. As long as ambient illumination is provided by a number of fluorescents, the phase difference will be large enough to eliminate this effect.

The precise character of the various effects was dependent upon the particular conditions of the electronic variables (the spatio-temporal distribution of photic energy as generated by the beam), the format of the display, and the illumination conditions. The three classes of phenomena are:

a. Jump

The term "jump" refers to the very rapid apparent displacement of display content when the viewer's eyes are moved. The symbols appear to jump and then return to their normal positions. The appearance of this jump is simple in the case where only one or very few symbols are present; the single symbol seems to jump from its normal position and then to return after a very short time interval. When the display is composed of a number of lines of symbols, the phenomenon becomes perceptually more complex. The structure of the format exerts a restraining influence on the jump so that parts of the display seem to move while other parts do not. This phenomenon is difficult to describe, partly because of its unfamiliar nature, and partly because its particulars are different according to the direction, length, and end-points of the (saccadic) movements of the viewer's eyes.

b. Half-Page Shift

This refers to the apparent relative motion that can occur between the top and bottom halves of a display format. The phenomenon is dependent upon there being a pause between the writing of the top and bottom halves. In contrast to the jump phenomenon, the half-page shift occurs when the observer gazes steadily at the display without moving his eyes. The most pronounced effect occurs between the bottom line of the top half and the top line of the bottom half. The symbols in these lines appear to undertake a random "dance" with respect to one another. The movement is not confined to one line moving with respect to the other, although this is a predominant motion, and most of the movement is lateral rather than horizontal. When the observer's eyes are moved, an apparent movement of all (or of some large area) of one half-page with respect to the other can also occur. This movement is similar to the jump movement described above, but it takes place with respect to large areas and is seen as a definite break between the top half and the bottom half.

c. Veiling

This refers to the perception of shifting patches of filmy light moving over the tube face. The phenomenon is associated with fairly dense formats of symbols. We believe it is due to variations in total light flux in different areas of the tube face. It is most pronounced when one views the tube face steadily and

is decreased when the eyes are moved around the display. The effect can merge into a more systematic sweep of light associated with the refreshing of the symbols. It can also merge with any flicker that may be perceived, most noticeably in the periphery of vision.

Before proceeding to a discussion of the phenomena described above, it is important to point out that their appearance and potency are dependent on a large number of conditions. The visual situation can be "tuned" to produce all the effects simultaneously and in their most potent forms. Similarly, they can be "tuned out." The simplest way to tune them out is by using a long-persistence phosphor. For instance, none of the phenomena described above could be produced using a P-12 phosphor. However, because a slow phosphor may impose certain penalties, it is considered desirable to arrive at some understanding of the boundary conditions of the phenomena. This will permit consideration and evaluation of the various options for ridding the displays of the movement phenomena.

3. Analysis of the Phenomena

a. The Jump Phenomenon

The jump phenomenon occurs only when the observer's eyes are in motion. When scanning freely (as opposed to tracking a moving target), the eyes move in a series of sudden motions, called saccades, which are interspersed with fixations.

Observation of the jump phenomenon shows that the direction of the jump seems to be opposite to the direction of motion of the eye. Hence, the basis of the phenomenon must be that successive pulses strike the rotating retina at different locations. For example, if the eyes should move from right to left, a stimulus would be seen first in the left visual field and then in the right visual field, and its apparent movement would be from left to right (that is, in the direction opposite to the eye movement). While there can be little doubt that this is the basic mechanism of the phenomenon, our confidence in its trueness is strengthened by the observation that the jump appears more pronounced when the saccade is made over large angles rather than small angles. The relevance of this observation is that peak angular velocities increase as saccade angle increases. Sample figures are:

Saccade Extent	Peak Velocity *
15°	320° per second
30°	540° per second
60°	680° per second
90°	720° per second

*Hyde, J.E. "Some Characteristics of Voluntary Human Ocular Movements in the Horizontal Plane." J. Amer. Opthal. 1959, 48, 85-94.

Hence it is to be expected that the amplitude of the apparent jump should be correlated with the angular distance the eye moves in the time elapsing between successive pulses.

It also seemed (at least during our observations) that when the proper conditions were in effect, the jump phenomenon always occurred during long saccades but sometimes did not occur during short saccades. This is consistent with the finding that the length (in angular distance) of a saccade is correlated with the time duration of the saccade. Average figures are:*

Saccade Extent	Time Duration
2^{o}	15 ms
5^{o}	29 ms
10^{o}	39 ms
15^{o}	48 ms
20^{o}	55 ms
30^{o}	80 ms
40^{o}	100 ms

Thus, when the refresh interval is, say, 25 ms, there will be a fair number of saccades (e.g., those used in reading text, which may average out at about 3^{o} or 4^{o} in extent) which will be completed in less time than the refresh interval.

To summarize, the jump phenomenon is thought to be due to successive pulses falling on different points of the retina; the magnitude of the phenomenon is related to the velocity of the angular rotation of the eye, which is highest for long movements; and the probability of the phenomenon's occurring is related to the duration of the eye movement, which is longest for long movements.

For the jump phenomenon to be seen, it is necessary for the eye to see while in motion. In spite of a mistaken notion, still widespread, that the eye goes blind during saccades (the apparent reason for our not seeing a blur between fixations while scanning), the fact is that the eye does see during saccades, but not very well. In terms of brightness threshold, to be seen during a saccade, a stimulus may have to be brightened by a factor of 10 (as compared to brightness required when the eye is still). Hence, in general, we can expect the jump phenomenon to be perceived most strongly when the stimulus is bright and when the stimulus/background contrast is high. Observation of different brightness conditions on the dd60 equipment seemed to bear out this expectation.

The format of the visual stimulus also seems to exert some influence on the appearance of the phenomenon. In general, our observations led us to the conclusion that a simple element appearing in isolation has the greatest

*Woodworth, R.S. and Schlosberg, H. Experimental Psychology (rev. ed.). Henry Holt and Company, New York, 1954.

tendency to jump and that large blocks of densely packed symbols have less tendency to jump. There seemed to be a competition in the latter case for the symbols to jump on the one hand and to remain still and cohere with the rest of the rest of the material on the other hand. However, on the whole, we felt that the factors tending to cause some appearance of jump could easily overwhelm the factors predisposing to stability, so that the format variables took on major significance only when the jump conditions were weak.

Of all the variables which enter the phenomenon, the strongest is undoubtedly the distribution of light over time. For the conditions of display on the dd60 equipment, jump could not be seen when refresh intervals were at or about 15 ms; with dense format, 20 ms was the minimum interval. Jump became very obvious at 25-ms intervals and continued to grow worse up to and through the CFF point. It is not clear at this time what factors are responsible for setting the threshold points. The P-31 phosphor decays to a 10% value in 40 μs and thus is short enough, one might think, for jump to be seen at intervals shorter than 14 or 20 ms. Perhaps the explanation lies in the fact that the angular extent of movement is small (2^{o} to 4^{o}) in these time intervals and that the eye has sufficient spatial integrating capability to avoid seeing the extents as jumps; however, this explanation is not very plausible when it is recalled that the visual acuity of the still eye is a few seconds of arc and the threshold for movement (in central regard) is one to two minutes of arc per second. Further analysis and experimentation is required before an adequate understanding of the phenomenon can be achieved.

b. The Half-Page Phenomenon

The nature of the phenomenon has already been described. Viewing the characters through a microscope (on the dd60 equipment) satisfied us that the apparent movement was not due to the objective movement of the symbols. The amplitude of such movement was not greater than about .5 ms; and in any event, the objective movement within a given half-page was constant, and no similar movement phenomena were seen elsewhere in that half page.

Of the various observations we made, the following are the most indicative of the conditions of the phenomenon:

1) The phenomenon persisted when the display was reduced to one character in the top half and one character in the bottom half; time interval between painting was about 10 ms.

2) The phenomenon grew less noticeable as the top and bottom lines were separated and was absent when they were an inch apart.

3) The apparent movements did not follow any set pattern and were characterized by a large degree of randomness in time and space.

While, in general, the phenomenon seems to be due to the fact of succession, which the eye interprets as movement, the details of the phenomenon do not conform to any published accounts which we have read. For one thing, the time interval is much shorter than the time interval previously reported to be conducive to the perception of apparent movement. Bartley* reports the optimum interval to be about 60 ms. However, neither he nor other workers has reported the threshold conditions for apparent movement. Some reports would lead one to believe that movement can be seen when stimulus A is turned off and at the same point in time stimulus B is turned on and when A and B are at separate but adjacent points. If one accepts this possibility, then apparent movement will typically be seen on a P-31 phosphor when the writing interval is greater than about 40 μs (the decay time to 10% of brightness) and when the brightness and juxtapositioning of the adjacent stimuli are conducive to the perception of movement.

The whole story, however, must be more complex than this, for it takes more than 100 ms to write each line; and, it is recalled, no movement phenomena comparable to the half-page shift occur within a half page. In other words, there must be some time interval value between about 1/10 ms and 7 ms (the shortest time interval at which we have observed the half-page shift) which brings about the half-page shift phenomenon.

There is a possibility that the observed effect is due to an interaction between time/energy characteristics of the illuminated characters and normal eye tremor. In a laboratory demonstration it was possible to observe two small neon lights appear to jump apart and then return at certain combinations of flash rates and phase relationships, but only during saccadic eye movements. Again, further analysis and investigation is required.

4. Discussion

The crux of the problem can be thought of in terms of the relative amounts of photic integration that are performed in the phosphor and in the eye. When the phosphor decay persists long enough so that only a small decrement in brightness occurs before the next refreshing signal, no movement phenomena occur. When this decrement in brightness is large, and when time intervals are suitable, the apparent movement phenomena appear.

*Bartley, S.H. The Principles of Perception. Harper and Row, New York, 1958

Selecting a longer persistence phosphor (see Section B, below) will eradicate the movement phenomena at the cost of some decrease in display brightness and some "smear" which occurs when the display is updated. Other possible solutions, in terms of extending the character writing time and/or of increasing the PRF rate, carry the penalties of reducing the amount of information that can be written.

As with so many problems of this type, one is faced with making tradeoffs. However, with this problem we faced the disadvantage of not knowing in quantitative terms either the conditions of display that produce the undesirable visual phenomena or their effect upon human operator performance in checkout tasks. Hence, we believed it imperative that performance degradation be assessed first and that the quantitative boundary conditions for the phenomena then be established.

B. An Experimental Investigation of Operator Performance Using P-31 and P-12 Phosphors with the dd60 Display Unit

The experimental work described below was undertaken to provide information on the effects of these phenomena on operator performance. Accordingly, two experiments were conducted to compare operator performance using a fast (P-31) phosphor and a slow (P-12) phosphor. The electronic parameters were so designed that the desired apparent movements were present when the P-31 phosphor was used and not present when the P-12 phosphor was used. Experiments I and II were similar except for the work-rest cycle, which was designed to increase operator loading during Experiment II. Both experiments were conducted at Data Display, Inc., using a dd60 display unit. Task and environmental conditions were similar to those encountered in the operational situation. A detailed account of the work and the preliminary results obtained is given below.

1. Method

The total task was comprised of a primary and a secondary task. The primary task required the subject to answer questions on a test card by using information presented partly in the CRT display format and partly on a permanent placard directly above the scope. Four test cards of twenty questions each were prepared, one for each of four display formats. Each format contained twenty-six lines of alphanumeric information, selected and organized to simulate the essential properties of a scope format during operational checkout. The permanent placard contained seven lines of information similar in content and organization to that in the scope formats. Figure 1 shows the placard and representative sections of a test card and of a scope format.

	Address	Sign	Code	Sender	Message
A	9P8320	-	19	JAV	RR STRBD. ROCK. EMPT.
B	53LZC	-	8		KL JAW EXPRESSWAY 953
C	629N10		025	TIS	ESC. VELOC 25 KM
D	32532Q	L	5	KO	ALEN FIN. LITE OK
E	JU8185	+	21	BIM	EEG BETA WGHT 24Z WGHT 21
F	59LP219	=	0	ARW	SHARD FREQ 90 KC
G	65210	-	AGC		SUPER HET. SINK SIGNAL -

PLACARD

12. Write Sender of Signal 6

13. Is the sign appearing in Placard Signal C the same as the sign appearing in Signal 5?

14. Which signal refers to BOATED?

15. Write out the part of the signal which comes before... ...B 9I

16. Add the digits (single numbers) appearing last in the addresses of Signals 20, 9, and 15

TEST CARD

93A661	+	3	GO	WIRE BLOOD RATE 66LB
91320B		9	ING	BOATED EFFECT 633 FT
22321B	+	3	TUB	CREWDOW CHGE + 63G
23322B	+	6	TH	FIL. FEEDTHRU 9 KV AT 1
33333B		4	EBE	CONTROLLER GRIP 6 LB
42C865	+	0	ST	PRINK ALARM 02. 00 HR
89D625	-	19	WU	MERCURY FALLOUT 1% K

FORMAT

Figure 1. Placard and Sample Sections of a Test Card and Scope Format

As shown in Figure 1, task material varied from meaningless groups of alphanumeric characters to words and statements having a contextual face validity. The questions were designed to require large saccadic eye movements among the test card, placard, and format and both large and small movements within the format. Also, questions ranged in difficulty from those requiring simple transcription, e.g., Question #12, to those requiring elemental quantitative skills, e.g., Question #16.

The secondary task, which was performed concurrently with the primary task, required the subject to monitor audio communications continuously. Inputs consisted of one digit or one of twenty common male first names, each having a unique first letter. The subject was instructed to record on his answer sheet the first letter of any name he detected.

In combination, these tasks presented the subject with a total task comparable in critical behavioral characteristics to actual checkout tasks as described by operational personnel. The first task required the appropriate scanning, reading, cognitive, and transcribing behavior; while the second task demanded concurrent attentiveness to critical audio inputs.

This combination of tasks was employed in both experiments, with the major differences between the experiments being task duration and the work-rest cycle. These specific differences are discussed in Section 5, on experimental procedures.

2. Equipment Description and Layout

The study was conducted in a curtained area of the equipment checkout room at DDI. Major environmental variables were controlled to the following limits:

. Noise	76 db ± 4 db
. Temperature	70° F ± 2° F
. Ambient Illumination	5 footcandles incident to plane of scope face

A digital computer was used to program the dd60 display unit. The unit consisted of two 12-inch CRT's, one coated with a P-12 phosphor, the other coated with a P-31 phosphor. Electronic parameters of the scope displays were selected to reproduce the essential characteristics of the dd74G presently in operational use. Because of limitations inherent in the computer simulation and the unavailabiltiy of a dd51 or -74G unit, faithful replication was not possible.

The major electronic characteristics of the dd60 and dd74G are compared below:

	dd74G	dd60
. pulse repetition rate	19.7 ms	19.7 ms
. character time	4.8 μs	7.2 - 7.4 μs
. flyback time	.8 μs	.8 μs
. half page print time (lines 1-13 or 14-26)	2.5 ms	6.1 ms
. transfer time - upper half to lower half	15.0 ms	7.4 ms

In terms of reproduction of the visual phenomena in question, the replication appeared to be adequate. DDI personnel and members of our own staff who had viewed the display units reported all phenomena under study were present on the dd60 unit

Illumination variables were precisely calibrated and monitored throughout the study. As indicated above, ambient illumination was 5 footcandles at the scope; background brightness at both scopes was 3 footLamberts, with daily maximum adjustment of 8%. Between-scope differences were minimal -- less than 5%. Some small change in brightness for other areas of the scopes could be expected. However, area by area, the scopes had identical brightness $\pm$ 5%.

In addition to the display unit, equipment within the enclosure included a console-type adjustable chair for the subject, a work table and chair for the experimenter, and a Revere table-model tape recorder used as the auditory stimulus generator for the secondary communications task.

3. Subjects

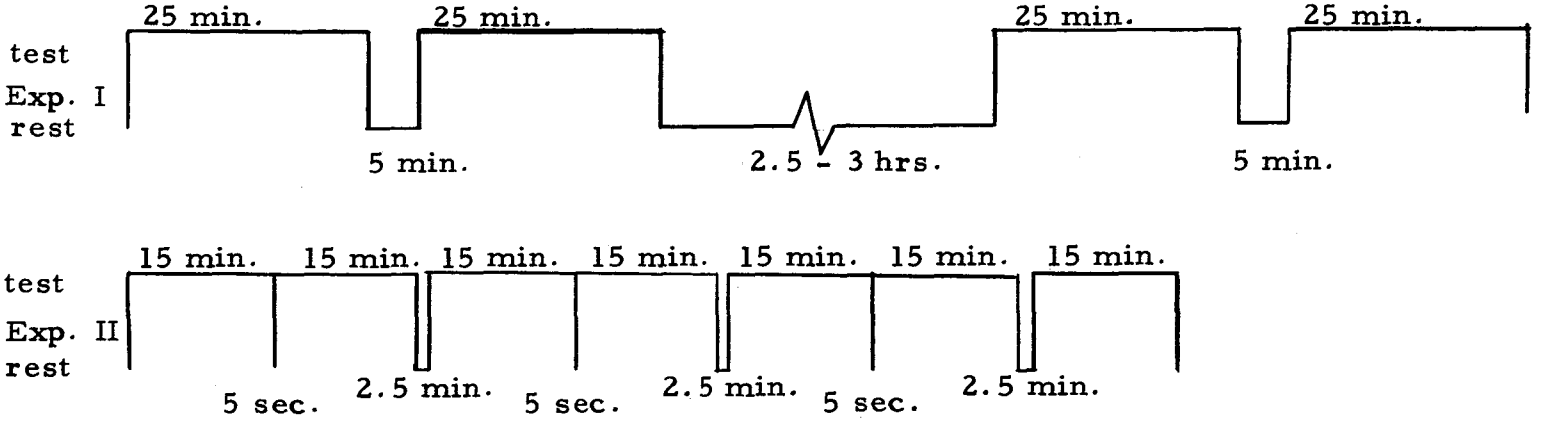
Eight subjects were used in the first experiment. Four of these subjects were also used in the second experiment. All subjects were male college graduates recruited through contact with local public schools and graduate schools of colleges in the area. Several subjects had advanced degrees or were matriculants for advanced degrees. Six subjects were public school teachers; two subjects were ministers. The average age of the eight was 32 years, with a range from 27 to 43 years. Five subjects had normal vision; three subjects had corrected normal vision.

4. Experimental Design

In both experiments, subjects were run as their own controls in a counterbalanced design. Repeated measurements were taken in Experiment I across the scope conditions. In Experiment II, each subject was given one extended trial on each scope condition, with the order of conditions counterbalanced across subjects. Other variables counterbalanced in both experiments were: format order, test card order, tape order (the communications task) and time of testing.

5. Procedure

Prior to experimentation, all subjects were given a three-hour group orientation to the study. Included in the orientation was a 15-minute demonstration trial on the task. In Experiment I, each subject was given four trials (two per day), with each trial consisting of two 25-minute tests. The rest period between trials was several hours; between tests it was 5 minutes. In Experiment II, each subject was given two trials (one per day), with each trial consisting of six 15-minute tests. Rest periods between tests averaged less than 2 minutes. A time-line plot of the work-rest cycle for one day in each experiment is shown below.



As indicated, the intent of Experiment II was to load the task by minimizing duration of rest in the work-rest cycle. Post-test questionnaires indicated the procedure was effective, in that all subjects reported the task in Experiment II to be more difficult, i.e., more fatiguing.

6. Measures

In each experiment, four performance measures were obtained on each test. One speed and one accuracy measure were obtained on the primary task (the 80-item test). Two accuracy measures (commission errors and omission errors) were obtained on the secondary task. In addition, a post-test questionnaire was administered at the completion of each experiment. The questionnaire was identical for both experiments, and the final analysis will include study of the questionnaire responses as an indication of differential subjective experiences in each experiment.

7. Results and Discussion

The results presented herein are based solely upon a preliminary "desk calculator" analysis of the raw data. A detailed computer-processed analysis of variance is currently being completed, the results of which will be presented and discussed in a special study report, which should be released during the next quarter.

The preliminary results obtained in Experiment I are shown in Table 1, which shows the averaged performance of all subjects on each test and for each phosphor. Also shown is the average on all tests for each phosphor against each performance measure. These are shown in two ways: (1) as the percentage performance improvement that was achieved by using the superior phosphor; and (2) as the percentage decrement associated with the poorer phosphor.

The performance measures are as discussed in Section 6, with the exception that errors of commission and omission on the secondary task are combined to give a total error score. As indicated in the table, all differences in performance for each phosphor are relatively small and are probably not statistically significant.

While final evaluation of the results must await detailed statistical analysis, our tentative conclusion is that operator performance does not vary substantially as a function of the different phosphors investigated under the conditions of Experiment I. It would appear that under these light task loads, the phenomenon has no degrading effect or that its effect is compensated for by some other aspect of the phosphors (e.g., the color difference between the P-31 and the P-12).

The results of Experiment II (higher task load) are being analyzed in a similar fashion, and preliminary results are shown in Table 2. The results indicate that speed on the primary task does not differ systematically with type of

Table 1

Preliminary Results from Experiment I

Test No.	Primary Task Items Attempted per Minute P-31	Primary Task Items Attempted per Minute P-12	Primary Task % Errors P-31	Primary Task % Errors P-12	Secondary Task % Total Errors P-31	Secondary Task % Total Errors P-12
1	1.93	2.06	8.83	7.57	3.09	2.02
2	2.55	2.52	9.62	7.29	2.04	3.18
3	2.98	2.50	7.56	10.00	1.54	3.02
4	2.87	2.83	4.24	6.89	1.65	4.28
5	3.14	2.89	7.63	7.79	.72	2.12
6	3.17	3.20	5.63	2.90	1.02	1.19
7	3.56	3.36	5.79	5.06	3.23	0.00
8	3.47	3.59	3.75	6.60	1.09	0.00
Average	2.96	2.87	6.63	6.76	1.80	1.98
% degradation using P-12 instead of P-31 phosphor		3.0		2.0		10.0
% improvement using P-31 instead of P-12 phosphor	3.1		1.9		9.1	

Table 2

Preliminary Results from Experiment II

Test No.	Primary Task Items Attempted per Minute P-31	Primary Task Items Attempted per Minute P-12	Primary Task % Errors P-31	Primary Task % Errors P-12	Secondary Task % Total Errors P-31	Secondary Task % Total Errors P-12
1	3.60	3.70	7.32	3.79	2.53	.89
2	3.93	3.85	2.35	3.72	.86	0.00
3	3.55	3.62	3.52	3.81	1.73	2.56
4	3.43	3.75	2.35	2.95	.81	.83
5	4.00	4.02	4.90	2.69	.89	.89
6	3.92	3.85	7.05	3.61	1.70	.83
Average	3.74	3.80	4.58	3.43	1.42	1.00
% improvement using P-12 instead of P-31 phosphor		1.6		25.1		29.6
% degradation using P-31 instead of P-12 phosphor	1.6		33.5		42.0	

phosphor. However, the results on accuracy measures for both tasks are sufficiently large to suggest that these differences possibly may be significant and of practical import. (This will, of course, have to be confirmed by more complete analyses of variance.) As shown in Table 2, when the P-31 phosphor is used, errors are 33.5% greater on the primary task and 42.0% greater on the secondary task.

If the forthcoming detailed statistical analysis substantiates these findings, then a justifiable summary conclusion would be: The apparent movement phenomena reported with the P-31 phosphor do not affect performance under relatively light task loading (as per Experiment I), but as task loading increases (as per Experiment II), these phenomena degrade the accuracy of performance both on the primary visual task and on the concurrent secondary audio monitoring task.

C. Planned Study of Apparent Movement Phenomena on Pulsed Displays

During the next quarter we will complete the analysis of the foregoing experiment and conduct a simple controlled laboratory investigation to examine the conditions for the appearance of apparent movement phenomena on pulsed displays. It has been established that certain combinations of frequency, pulse length, brightness, and other factors can give rise to displays which do not flicker but which appear spatially unstable. As noted earlier, the two main effects appear to be a jump phenomenon (the rapid apparent translation of the stimulus source to other locations) and a shift phenomenon (a continuing shifting of one display element with respect to another). These phenomena are provisionally attributed to stroboscopic effects associated with saccadic eye movements for the jump phenomenon and with eye tremor for the shift phenomenon.

The experiments are designed to determine the boundary conditions for the appearance of the phenomena under laboratory conditions. These conditions can then be used as a basis for preparing design specifications for the manner of writing on short-persistence CRT's in order to avoid the appearance of spatial instability phenomena.

To reduce expenses and at the same time gain precise experimental control of the variables involved, we will use experimental equipment consisting of optical paths interrupted by rotating sector wheels which will display symbols on a screen. To investigate the jump phenomenon, a single optical path will be employed to project a symbol or symbols, and the following will be varied:

- Frequency

- Pulse Length
- Brightness
- Ambient Illumination
- Size (of image)

To investigate the shift phenomenon, twin optical paths will be employed to project symbols in adjacent positions, and the following will be varied:

- Frequency
- Pulse Length
- Phase Relationship (between the two optical paths)
- Brightness
- Ambient Illumination
- Sizes (of images)
- Adjacency of Images

Subject observers will be drawn from D&A staff, and the psychophysical "method of adjustment" will be used to map the boundary conditions for the appearance of the phenomena.

PROGRAM ADMINISTRATION

A. Staffing

The following table indicates the personnel assigned to this project during the past quarter, the number of hours expended by each, and the tasks on which they worked.

Name	Hours	Tasks
Affinito, Frank	261	Visual Experiment
Bowen, Dr. Hugh	139	Visual Experiment
Eckenrode, Richard	409	II
Grant, George	169	All (deputy to Program Director)
Gruber, Dr. Alin	10	Visual Experiment
Guinness, G. Victor	368	Visual Experiment
Kahn, Edward	288	IV
Koller, Dr. Alice	220	All
Matteis, Richard	148	III
McCay, Robert	5	Visual Experiment
Nichols, Hugh	80	V
Pepler, Dr. Richard	164	II, V
Shah, Vishwa	271	IV
Stowens, Bernard	164	I, III
Tolcott, Dr. Martin	2	All
Wohl, Joseph	251	All (Program Director)
Technical Assistants	501	All

Although the hours may fluctuate from quarter to quarter as the emphasis on various tasks changes, these people form the core of the project staff.

As indicated in the introduction, the effort was not uniform across all tasks. The following table shows the approximate man-hours devoted to each task.